Dirac's Quantum Electrodynamics

Alexei Kojevnikov

Dirac's relationship with quantum electrodynamics was not an easy one. On the one hand, the theory owes to him much more than to anybody else, especially if one considers the years crucial for its emergence, the late 1920s and early 1930s, when practically all its main concepts, except for that of renormalization, were developed. After this period Dirac also wrote a number of important papers, specifically, on indefinite metrics and quantum dynamics with constraints. On the other hand, since the early 1930s he was an active critic of the theory and tried to develop alternative schemes. He did not become satisfied with the later method of renormalization and regarded it as a mathematical trick rather than a fundamental solution, and died unreconciled with what, to a large extent, was his own brainchild.[1]

In the present paper I examine Dirac's contribution to quantum electrodynamics during the years 1926 to 1933, paying attention to the importance and the specificity of his approach and also tracing the roots of his dissatisfaction with the theory, which goes back to the same time and which, as I see it, in many ways influenced his attitude to its subsequent development. Some of Dirac's crucial accomplishments of that period, in particular his theory of the relativistic electron, have already been studied by historians in much detail. I will describe them more briefly, placing them in the context of Dirac's other works and of the general situation in quantum theory and leaving more room for other, less studied works, such as the 1932 Dirac–Fock–Podolsky theory.

1. 1926: The Compton Effect

This was the year when quantum mechanics was still being created and, at the same time, attempts were already being made to develop a relativistic quantum theory and to study problems related to radiation. It was not yet

Yuri Balashov and Vladimir Vizgin, eds., *Einstein Studies in Russia. Einstein Studies, Vol. 10.* Boston, Basel, Birkhäuser, pp. 229–259.

clear whether all this would lead to one unified quantum theory or, as it eventually happened, would split into two separate disciplines. Dirac was at the center of all these developments.[2]

On the subjects that would later become part of quantum electrodynamics, he wrote two papers containing the theory of the Compton effect. The discovery of that effect in 1923 led to the triumph of Einstein's light quanta and provoked a crisis of the old quantum theory. But that was just the beginning rather than the end of the history of the problem.[3] In the experimental studies prior to 1926, the formula for frequency change and the validity of conservation laws were reliably established, but the formula for the intensity of scattered radiation (i.e., the scattering cross section) and its dependence on the angle and frequency was not yet fully clarified. Theoreticians were proposing competing models of the effect, which continued to be one of the most actively debated problems in physics.

In April 1926, Dirac completed a paper that became, in the framework of new quantum mechanics, the first attempt at a relativistic generalization of the theory. Dirac was not yet employing Schrödinger's formalism there but followed his own, the so-called algebra of *q* numbers, which was a generalization of the Göttingen matrix mechanics.[4] According to Dirac's analytical approach, the transition to a relativistic theory had to be achieved by getting rid of the special treatment of the time variable t and by making it appear in the formalism in a way similar to space variables x, y, z. He described the electron according to quantum mechanics with relativistic corrections, while radiation was treated classically as an electromagnetic wave. Dirac considered a system consisting of an electron and an incident wave, and calculated the emitted radiation with the help of the conventional formula of matrix mechanics, obtaining through this procedure expressions for the frequency and the intensity of scattered light (Dirac 1926a).

His formula for the intensity,

$$I = \frac{I_0}{R^2}\left(\frac{e^2}{mc^2}\right)^2\left(\frac{\nu'}{\nu}\right)^3 \sin^2\theta, \tag{1}$$

where ν, I_0 are the frequency and the intensity of incident and ν', I of scattered radiation, R the radius and θ the angle of scattering, was only approximately correct. The exact formula would be derived by Oskar Klein and Yoshio Nishina by the end of 1928 on the basis of the 1928 Dirac relativistic wave equation for the electron with spin. All previous attempts (apart from Dirac, the problem was also attacked by Klein and by Walter

Gordon) provided only approximate results, though quite close to the final one (Gordon 1926, Klein 1927, Klein and Nishina 1929).

Dirac's paper was repeatedly used and referred to (see Small 1983), but his approach was not in itself popular. Practically all physicists who dealt, in 1926 and early 1927, with the relativistic problem in quantum theory treated it by the methods of Schrödinger's wave mechanics rather than matrix mechanics. Solutions for the electrodynamical problems were looked for on the basis of the Klein–Gordon equation combined with the interpretation of the wave function as the density of the electromagnetic charge and with the semi-classical description of radiation. Dirac, on his part, viewed this approach with skepticism. He accepted wave mechanics as a mathematical method only, without its physical interpretation, and that is how he used it in his second paper on the Compton effect in November 1926. It practically reproduced the approach and results of the first paper, only that a number of calculations were drastically simplified thanks to the use of Schrödinger's formalism (Dirac 1927a).

2. 1927: The Quantum Theory of Radiation

The first completed formalism of quantum mechanics—with its basic equations, methods of their solution, and rules of how to compare results of calculations with experiment—was developed very quickly after Max Born's proposal of the statistical interpretation of the wave function. Dirac, and simultaneously and independently Pascual Jordan, accomplished this in November 1926 (Dirac 1927b, Jordan 1927a,b) and soon declared quantum mechanics done.[5] "The new quantum theory . . . has by now been developed sufficiently to form a fairly complete theory of dynamics," Dirac wrote in February 1927, adding: "On the other hand, hardly anything has been done up to the present on quantum electrodynamics" (Dirac 1927c, p. 243).

The quotation reflects both Dirac's refusal to accept the electrodynamics of Klein and Gordon and the conclusion that quantum mechanics had developed as an essentially non-relativistic theory. Dirac's formulation of its basic principles was grounded in the Hamiltonian method, in which time had a special role as a parameter and could not be treated symmetrically with space variables. Therefore, according to Dirac, there was a need for a new separate theory of quantum electrodynamics that would account for relativistic problems as well as for a consistent description of the electromagnetic field.

In the same paper, Dirac made a partial step towards this future theory, proposing a quantum theory of electromagnetic radiation and its interaction with matter. In later accounts, this paper is often referred to as "the beginning" of quantum electrodynamics, although, of course, one cannot say that quantum mechanics had never up to that point considered radiation. A number of phenomena could be accounted for by describing radiation according to the classical wave theory and considering its interaction with a quantum-mechanical particle. These methods were applied to calculate the effects of the scattering of light by an electron (dispersion and the Compton effect) and of atomic quantum transitions caused by incident radiation. In addressing the latter problem, several authors, including Dirac (Born 1926, Dirac 1926b, Slater 1927), had been able to calculate the value of the Einstein coefficient for stimulated emission and absorption of radiation (the so-called *B* coefficient), but not the other, *A* coefficient for the probability of spontaneous emission.

To obtain the *A* coefficients, which determined the intensities of spectral lines, quantum mechanics used an additional rule, called Heisenberg's hypothesis, which was, in fact, the very first postulate from which quantum mechanics had started in 1925. It asserted that the probability of radiative transition is proportional to the square of the corresponding matrix element of the coordinate of the electron. Dirac eliminated the need for an additional postulate, having derived both coefficients together based on a single approach, in which waves of radiation themselves were described quantum-mechanically.[6] Each harmonical component of the radiation was quantized as an oscillator according to the rules of quantum mechanics. This treatment had already been proposed by Jordan in late 1925, in one of the first papers on matrix mechanics (Born, Heisenberg, and Jordan 1926), but only gained wide recognition after Dirac's effective demonstration of its power in 1927. Dirac's approach was instantly welcomed as the first consistent quantum theory of radiation and accepted as the paradigm in a whole series of subsequent studies.

Although very successful in this practical sense and capable of describing an ever increasing scope of phenomena, the theory, by Dirac's own criteria, still lacked a lot to become a consistent quantum electrodynamics. Since radiation was quantized as a Hamiltonian system, the relativistic invariance of the theory was not apparent. Secondly, the quantization dealt only with the radiation part of the electromagnetic field without the Coulomb part for the interaction of charges, and thirdly, the particles themselves were described as non-relativistic. Dirac saw the first of these handicaps as an especially serious one and, because of it, did not consider his accomplishment as the starting point of a future consistent theory, the present appraisal notwithstanding. The latter, he believed, had to be relativistically invariant right from the start and explicitly (Dirac 1927c, p. 243–4).

Things actually developed differently, by way of gradual corrections of the existing shortcomings. In 1928 Jordan and Wolfgang Pauli proved the relativistic invariance of the quantization of radiation, while Dirac found the relativistic wave equation of the electron. The following year, Enrico Fermi and also Werner Heisenberg together with Pauli extended the method of quantization to the full electromagnetic field, together with its Coulomb part (Jordan and Pauli 1928, Dirac 1928a, Fermi 1929, 1930, Heisenberg and Pauli 1929, 1930).[7]

3. 1927: Second Quantization

Apart from the above method of wave quantization, Dirac's 1927 theory contained two other ways of describing radiation. Radiation is treated, in both of them, from a corpuscular perspective as an ensemble of photons, quantum-mechanical Bose particles in the relativistic limit of zero mass and the velocity of light. Dirac derived the wave equation for an ensemble of bosons by two different methods: by imposing the requirement that the wave-functions of many-body systems must be symmetrical and by quantizing the wave function of a single particle, or, in modern terminology, through the second quantization.

This somewhat bizarre term stands for an idea that was also considered bizarre by many: to take the Ψ function of an already quantized system and make an operator out of it, that is, actually, to quantize the system for the second time. With the help of textual analysis, in particular, by paying attention to the evolution of Dirac's system of notation, one can reconstruct the history of the 1927 theory of radiation as originating from the attempt to quantize the wave function.[8] Dirac recalled later that he had not foreseen what would result from it and was sincerely surprised to find out that the procedure transformed the wave equation for one particle into an equation for a system of many Bose particles (Dirac 1983, p. 48). To confirm this unexpected result, he derived it once again by the conventional method of symmetrizing wave functions, meanwhile also obtaining, for the first time in the quantum theory of many-body systems, the wave equation for a Bose ensemble in the external field.

This result suggested to him applying the theory to photons and to their interaction with the atom. Pursuing the theory of quantum-mechanical Bose particles further, and considering its relativistic limit, Dirac managed to calculate the ratio between the two coefficients, *A* and *B*, but in order to obtain their absolute values, he had to supplement the photon theory with the method of quantized electromagnetic waves explained above. Both approaches provided results that were in good agreement, but the model of photons did not appear as mathematically powerful as the formalism of quantized waves.

Yet it was, of all the three Dirac methods of describing radiation, the one that satisfied his physical worldview the best, and he invested some more effort into it. In the first edition of *The Principles of Quantum Mechanics*, Dirac presented his quantum theory of radiation as the theory of photons, quantum-mechanical particles described by symmetrical Ψ functions and taken in the limiting case of relativistic velocities (Dirac 1930b). By that time, Dirac managed to develop the mathematics of the model enough to obtain both coefficients, *A* and *B*, and the dispersion formula without recourse to the formalism of quantized electromagnetic waves. He also abandoned the method of second quantization immediately after it had served its initial heuristic role. Dirac did not use it in his subsequent papers on quantum electrodynamics and on the many-body theory, leaving it to others to develop the approach further and recognizing it again only in the second edition of *The Principles* (Dirac 1935).

The method of second quantization was picked up by Jordan and then by Vladimir Fock (Darrigol 1986, Kojevnikov 1988b). Jordan's major achievement was to find out how to generalize the quantization of the wave function so as to obtain a system of Fermi particles (Jordan 1927c, Jordan and Wigner 1928). This gave him an opportunity to formulate the program of quantum electrodynamics on the basis of the fundamental concept of quantized waves applied to describe both the electromagnetic field and material particles, the program that has been realized in modern quantum field theory. Fock proposed a special representation (the Fock space) that allowed the translation of the formalism of second quantization at any stage into the language of conventional quantum mechanics, which removed the appearance of strangeness and became important for the method's general acceptance (Fock 1932).

Many modern presentations interpret second quantization wider, as the quantization of waves of any sort, including the electromagnetic waves. From this point of view, the first application of the method should be attributed to Jordan (Born, Heisenberg, and Jordan 1926), while Dirac is credited with its independent invention in the case of material particles. The mathematical procedure, indeed, is very similar in both cases, but it should be noted that this view reflects the modern perspective on second quantization. During the 1920s and 1930s, the quantization of the wave function, or of the matter wave, was viewed as a separate idea distinct from, and much more controversial than, the quantization of radiation waves.

4. 1928: The Dirac Equation

The many difficulties that preceded the development of quantum mechanics were later found to be caused by two separate problems: the electron's wave properties and its spin. Their combination only increased the confusion. Thus Erwin Schrödinger, for example, initially derived his famous equation in a relativistic form, which he immediately rejected because it gave the wrong spectrum for hydrogen. The non-relativistic wave equation proved to be a better approximation, while the relativistic wave was later found to provide good results only if spin was simultaneously taken into account.

In the crucial winter of 1925–26, both issues were clarified independently: George Uhlenbeck and Samuel Goudsmith proposed a visual model of the spinning electron, while Schrödinger published his wave equation. Its relativistic generalization, the Klein–Gordon equation, was quickly suggested by a number of authors (see Kragh 1984), and it was realized that the next step would have to combine both new results together into a quantum mechanical wave equation for the spinning electron. This logical path was taken by many: in 1927 Charles G. Darwin and Pauli solved the problem for the non-relativistic electron (Darwin 1927, Pauli 1927), and soon afterwards Hendrik Kramers, Jordan, Eugene Wigner, Yakov Frenkel, Dmitri Ivanenko, and Lev Landau examined the relativistic case (Kragh 1981b, pp. 61–62). Before any of these attempts succeeded, Dirac arrived at the result in a different way: he was preoccupied with creating a consistent relativistic quantum mechanics rather than with describing spin.[9] Spin came as an extra gift out of his equation, hence Dirac managed not only to unify spin with the wave mechanics, but, in a certain sense, to explain it (Dirac 1928a,b).

Since drafts or archive materials did not survive, historical reconstructions of this landmark achievement are based mainly on Dirac's published papers and his later reminiscences. Opinions differ as to whether he had no intention whatsoever of describing spin (as he himself claimed) or bore it somewhat in mind when he chose to play with the Pauli spin matrices. One way or the other, his primary motivation was his dissatisfaction with the Klein–Gordon relativistic equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + \frac{m^2c^2}{\hbar^2}\right)\Psi = 0, \qquad (2)$$

which he considered inconsistent with the basic core of quantum mechanics, the transformation theory.

His objections stemmed from requirements that, as later developments showed, could not be met. This led Dirac to admit, years later, that "the development of the relativistic theory of the electron can now be regarded as an example of how erroneous arguments sometimes lead to a valuable result" (Dirac 1959, p. 32). He presented these arguments in most detail in a little known paper in the summer of 1928 (Dirac 1928c). His objective was to find a wave equation with the always positive probability density. This could be met only for the density defined as $\rho = \Psi^*\Psi$, from which Dirac derived that the desired wave equation should be a differential equation of the first order in time. The equation (2) is of the second order, and it has an entirely different expression for the probability density.[10]

Looking for an equation which would be both linear in time and relativistically invariant, Dirac found it in the form

$$\left(\frac{1}{c}\frac{\partial}{\partial t} + \alpha_1\frac{\partial}{\partial x} + \alpha_2\frac{\partial}{\partial y} + \alpha_3\frac{\partial}{\partial z} + \frac{imc}{\hbar}\beta\right)\Psi = 0, \qquad (3)$$

where α_1, α_2, α_3, and β are 4×4 matrices of a special kind, and Ψ is, correspondingly, a 4-component wave function. Dirac further showed that, besides giving the positive probability density, equation (3) describes a particle with spin ½ and, in the first approximation, yields the correct formula for the fine structure of the hydrogen spectrum.

The publication of the Dirac equation had an immediate and tumultuous response. For more than a year, it attracted the attention and absorbed practically all efforts of physicists and mathematicians who dealt with quantum electrodynamics and relativistic quantum theory. Among its subsequent most important applications was the derivation of the final formula for the Compton scattering, the Klein–Nishina formula (Klein and Nishina 1929). The initial euphoria, however, soon gave way to a more sober attitude, especially since the fundamental difficulty of Dirac's theory was mentioned by him explicitly in the very first publication.

Dirac formulated it at first as an additional argument against the Klein–Gordon equation (Dirac 1928a), but he immediately saw that it remained unresolved also in his own theory. The relativistic equation had twice as many degrees of freedom as would have been sufficient for a particle with spin (four components of Ψ in the Dirac equation instead of two) which led to additional solutions corresponding to the states with negative values of energy. Similar solutions also appeared in the classical relativistic theory of a particle, but there they could be put aside and declared non-physical. The quantum theory could not get rid of them as easily, since the formalism did not forbid a transition of the particle into a negative energy state. The situation was further aggravated with the discovery of the "Klein paradox": electrons with positive energy could transform into negative energy electrons when passing through a potential barrier (Klein 1929). It was not clear how to interpret these theoretical electrons. The Dirac equation had great advantages and an enormous potential for application but also such serious problems that it was possible neither to discard it nor to put up with it.

5. 1929: Expectations and Disillusionment

While Dirac chose the dynamics of one particle as his way towards the relativistic generalization of quantum theory, his German colleagues made an attempt to create quantum electrodynamics as a many-body theory from the start. They treated both electromagnetic field and material particles as quantized fields described by Maxwell's equations in the former case and by the Schrödinger–de Broglie matter waves in the latter case. Jordan had formulated the basic principles of this approach in autumn of 1927 (Jordan 1927c), and he, together with Klein, Wigner, and Pauli, took the first important steps towards its realization (Jordan and Klein 1927, Jordan and Wigner 1928, Jordan and Pauli 1928).[11]

Early in 1928 Heisenberg and Pauli attempted to bring this program to completion.[12] At that time they still described particles with the help of the Klein–Gordon wave equation, which they wanted to quantize either according to Dirac (Bose particles) or according to Jordan (Fermi particles). But the general method of field quantization failed during the attempt to extend it from the electromagnetic radiation to the full electromagnetic field with the Coulomb part. Their collaborative work came to a temporary halt and resumed one year later, after Heisenberg had come up with a special mathematical trick to circumvent this difficulty.[13] An unusually long paper, "On the Quantum Dynamics of Wave Fields," was completed in March 1929 (Heisenberg and Pauli 1929).

The Heisenberg–Pauli electrodynamics brought all the earlier achievements—Dirac's radiation theory, second quantization, the Dirac equation—together in a comprehensive scheme and proved its relativistic invariance. But it did not live up to the high level of expectations. In the course of previous years, theoretical physicists had gotten used to quick and extraordinary success in sorting out the difficulties of the old quantum theory. In the case of quantum electrodynamics, however, problems of partial theories did not disappear with their unification into a general scheme; on the contrary, they became even more aggravated. All the difficulties of Dirac's electron theory remained unresolved, Heisenberg's "trick" with the so called ε-term was too artificial for a fundamental method of field quantization. Furthermore, hopes, initially raised by the partial Jordan–Klein theory, to subtract the infinite energy of the point-like electron by switching the order of quantum operators, were not realized in the more general treatment (Jordan and Klein 1927). This was only the first of many other infinities in relativistic quantum theory which would start attracting the attention of theoretical physicists in the subsequent years.

Disillusionment was not long to come. The authors themselves, Jordan first, followed by Pauli, Heisenberg, and by their close collaborators Ivar Waller, Robert Oppenheimer, and Leon Rosenfeld, were quick to announce that something was fundamentally wrong with the basic approach and that new radical changes were required. Such a pattern of behavior was not unprecedented: it also happened during the so-called crisis of the old quantum theory, which preceded the creation of the new quantum mechanics, though it was then confined to a narrower circle of participants and to a largely informal discussion. A similarly critical attitude to quantum electrodynamics became more widespread and more openly pronounced. Expressed readiness to give up some of its basic principles was characteristic of the crisis that lasted from early 1930 to early 1933. Exactly which radical changes to demand was not clear; Pauli, for example, expected the new theory to explain the value of the dimensionless constant $e^2/\hbar c$ (Pauli 1933).[14]

At first, Dirac did not participate in these developments. He wrote nothing on quantum electrodynamics in 1929 besides a section in his textbook *The Principles of Quantum Mechanics* where he presented his previous results. He did not include the Heisenberg–Pauli theory there, only mentioned it once in passing and in rather neutral terms (Dirac 1930b). It was not very likely that he was impressed by it, not only because of its problems, but also because of the underlying program which he did not share. Dirac still preferred the corpuscular version to the quantized waves approach in the treatment of radiation and matter. He also still tried to avoid using second quantization, in particular, in the paper of the same year dealing with the many-body problem in non-relativistic quantum mechanics

(Dirac 1929). But until 1932, he did not express his criticism of the Heisenberg–Pauli theory either.

6. 1930: The Hole Theory

While continental colleagues continued to struggle with quantum field theory, Dirac in Britain kept working on his own, somewhat narrower, topic, the relativistic mechanics of the electron. At the very end of 1929 he took a new important step in it. If one could not get rid of the troubling negative energy solutions of the Dirac equation entirely, it was still possible to try to ban the electron's transition to these states. This would happen, according to quantum statistics, if those states were already occupied with electrons. Dirac suggested that the negative energy solutions had physical meaning, but that the normal state of physical vacuum was the unobservable "sea" consisting of an infinite number of electrons occupying all possible states with negative energies.

Individual electrons with positive energies moved upon the face of the waters and were prohibited from falling into it by the Fermi statistics, because all positions below were occupied. If, however, a state of negative energy was free, there was a "hole" in the sea, which behaved as if it were a particle with normal, positive energy, but with the opposite sign of the electric charge. Dirac was tempted to identify this particle with the proton. His theory, then, would become a unified theory of matter covering both kinds of fundamental particles known by that time. Of course, the proton mass is almost 2000 times larger than the electron mass, but since the hole moves in the medium of negative energy electrons, Dirac maintained a hope to be able to derive the additional mass from the interaction between the hole and the sea. Dirac made this set of ideas public in a letter to Niels Bohr of 26 November and in lectures at the Institut Henri Poincaré in Paris in December 1929. He published them in a paper in January 1930 (Dirac 1930a).[15]

An electron of positive energy may fall into a hole, which would be observed as a mutual elimination, annihilation of both particles; and vice versa, if an unobservable negative energy electron gets a quantum of energy enough to jump to a positive state, a pair of an observable electron and a hole is created. Dirac's theory thus described the possibility of mutual creation and annihilation of material particles. The idea itself was not entirely new, since it had been discussed in astrophysics for several years by James Jeans and Arthur Eddington, and had also been mentioned by Jordan, but Dirac's model made it possible to calculate the probability of the process (Bromberg 1976, Jordan 1928). The opposition to Dirac's proposal was chiefly concerned not with this idea, but with the metaphysical, as many physicists thought, concept of the unobserved sea of negative energy electrons with its infinite density of charge and mass.

The response to Dirac's hole theory was as cold as the earlier reception of his electron equation had been enthusiastic. Pauli and Bohr authoritatively disapproved of it. Only a handful of quantum theoreticians supported Dirac: Igor Tamm welcomed the proposal without reservations, Oppenheimer accepted the idea of "sea," but not the identification of the hole with the proton. In his opinion, all holes had to be filled, since their annihilation with electrons occurs very rapidly, and one had to introduce another "sea" for protons. Tamm and Dirac separately calculated the probability of annihilation and realized that it, indeed, allowed the hydrogen atom to live only about 10^{-3} sec (Oppenheimer 1930, Tamm 1930, Dirac 1930c). Hermann Weyl studied mathematical transformations of the theory and became convinced that the hole mass had to be exactly equal to the electron mass; he even mentioned a positively charged electron, but only to say that it was not observed in nature and to return in the physical interpretation of the theory to Dirac's proton hypothesis.[16]

The difficulties of the hole theory or, rather, of the identification of the hole with the proton, were increasing: the mass difference could not be explained: on the contrary, there were hefty arguments in favor of mass equality; the hydrogen atom was stable and did not want to self-annihilate within a split second. All this caused Dirac to reconsider his proposal the following year.

7. 1931: Monopole and Other Particles

An interesting shift occurred in fundamental theoretical physics around 1931. Up to that point, the list of basic ontological entities of the world was likely to be seen as very short, usually consisting of three objects: gravitation, electromagnetic field, and the electron, and occasionally having the fourth one, the proton. And though it had already become apparent that some new forces were acting in the nucleus, and though physicists close to the experiment were occasionally discussing a hypothetical neutral particle, the high theory did not pay any serious attention to that. Much has been written about the great increase in the number of known elementary particles and interactions starting with the "miraculous year" of 1932, which saw the discovery of the neutron and positron. What is more interesting, however, is that the change of prevailing attitudes among theoretical physicists had begun even before those experimental discoveries.

In December 1930 Pauli, in a letter to a conference, put forward an idea that there might be a neutral particle in the nucleus. He needed it in order to solve two difficulties: with the statistics of nuclei and with the continuous spectrum of the β-decay. As would be understood later, these difficulties were to be ascribed to two different particles, a heavy one, the neutron, and an extremely light or massless one, the neutrino. Pauli envisioned the neutral particle to have spin ½, the magnetic moment, and a small mass comparable to that of the electron. He proposed a wave equation for it which was similar to Dirac's equation for the electron.[17]

In May 1931 Dirac submitted a paper that contained a theory of another hypothetical particle, the magnetic monopole (Dirac 1931a). He demonstrated that the idea of magnetic charge, which makes the Maxwell equations fully symmetrical, does not contradict quantum mechanics if the values of charges are connected by the relationship $eg = \frac{1}{2} n\hbar c$, where e is the electric charge, g the magnetic charge, and n an integer number. The monopole, if existing, would thus explain the fact of the quantization of electric charges. Although monopoles did not occur in experiment, Dirac sounded optimistic: "This new development requires no change whatever in the formalism. . . . Under these circumstances one would be surprised if Nature had made no use of it" (Dirac 1931a, p. 71).

Unlike Pauli, who was motivated by experimental difficulties, Dirac apparently came to his idea on the basis of purely theoretical speculation, hoping to explain the quantization of electric charge.[18] His predictions, however, did not stop with the monopole. In the same paper Dirac discussed two more unknown particles. Referring to the difficulties that arose in his hole theory with the proton mass and with annihilation rate, he abandoned the idea that holes were protons, suggesting instead that the theory calls for the existence of a light positively charged particle, the "anti-electron." Likewise, the proton would then require an anti-particle for itself (Kobzarev 1990, Kragh 1990). A real festival of new particles took place on 1 October 1931 in Princeton where both Pauli and Dirac presented reports on their recent proposals.[19]

Dirac was not absolutely sure in his prediction of the anti-electron but rather formulated the dilemma: either the new particle existed, or his electron theory had to be rejected, a possibility which he did not rule out totally.[20] The fact that the antielectron was not observed in experiment jeopardized his entire approach.

8. 1932: A Failed Revolution or How Dirac Nearly Became a New Heisenberg

The wave of critical attitudes towards quantum electrodynamics that was spreading from Germany reached Dirac in 1932, when he joined the work of critical reassessment of existing methods. At this point, his attention shifted from the theory of the electron to the quantum description of the electromagnetic field. The solution of the problem in Dirac's radiation theory (Dirac 1927c) and its further generalizations by Heisenberg, Pauli, and Fermi (Heisenberg and Pauli 1929, Fermi 1929) is considered to be correct now, but in the situation of crisis in the early 1930s, even those results aroused doubts.

In February 1931, Heisenberg proposed a new approach to electrodynamical problems (Heisenberg 1931). Unlike Dirac's radiation theory, it did not make use of the Hamiltonian function but relied directly on the equations of motion for the field and particles. A quantized electromagnetic wave excited the atomic system; the resulting charge and current densities were calculated using the rules of wave mechanics and determined, according to the classical formulas, the radiation emitted by the system. The theory thus obtained was close to the electrodynamical theory of Klein put forward in the earlier days of quantum mechanics (Klein 1927), with the main difference that incident electromagnetic waves were quantized rather than treated classically. Heisenberg offered his proposal as an alternative to Dirac's 1927 theory of radiation, arguing that a stricter reliance on the correspondence principle would pave the way out of the existing difficulties in quantum electrodynamics.

In a series of papers in 1931, Rosenfeld started developing this approach further and opposing it much more openly to Dirac's radiation theory (Rosenfeld 1931a,b,c). The latter, in his view, was responsible for the infinite values of the results of various calculations in quantum electrodynamics. Within the Heisenberg approach, Rosenfeld was able to reproduce all the basic achievements of the Dirac radiation theory and even to advance it somewhat further by deriving Christian Møller's formula for the scattering of two electrons with the relativistic retardation of the interaction (Møller 1931). It was probably Rosenfeld's critique that drew Dirac's attention to Møller's paper once he returned to Britain from a trip to the U.S. in early 1932.

By the standards of quantum electrodynamics, Møller did not derive his formula rigorously but guessed, to a certain degree, the correct answer. He considered the scattering of one electron on another in the Born approximation and, using a procedure similar to the semi-classical Klein–Gordon electrodynamics, established a correspondence between the electron's transition from one state to another and certain classical expressions for the densities of electric charge and current. The electromagnetic field thus produced, which he did not quantize but treated classically, acted on the second electron inducing its quantum transition into a new state. Despite the non-rigorous and non-symmetrical derivation, Møller's final formula for the scattering looked very reliable: it was symmetrical with regard to both electrons, relativistically invariant, and, indeed, was later fully confirmed within the fundamental theory. For the matrix element of the transition, in which the initial states of the electrons are set by the variables $\mathbf{p}_0{}^{\mathrm{I}}$, $u_0{}^{\mathrm{I}}$, $\mathbf{p}_0{}^{\mathrm{II}}$, $u_0{}^{\mathrm{II}}$ (u is the spin part of the electron wave function, $\mathbf{p}$ is its momentum), and the final states by the variables $\mathbf{p}^{\mathrm{I}}$, u^{I}, $\mathbf{p}^{\mathrm{II}}$, u^{II}, he obtained the expression

$$\Phi = \frac{e^{\mathrm{I}} e^{\mathrm{II}} \hbar^5 c^2}{\pi} \frac{< u^{\mathrm{II}} u^{\mathrm{I}} | 1 - \boldsymbol{\alpha}^{\mathrm{I}} \boldsymbol{\alpha}^{\mathrm{II}} u_0^{\mathrm{II}} u_0^{\mathrm{I}} >}{c^2 (\mathbf{p}_0^{\mathrm{I}} - \mathbf{p}^{\mathrm{I}})^2 - (E^{\mathrm{I}} - E_0^{\mathrm{I}})^2} \, \delta(\mathbf{p}^{\mathrm{I}} + \mathbf{p}^{\mathrm{II}} - \mathbf{p}_0^{\mathrm{I}} - \mathbf{p}_0^{\mathrm{II}}), \qquad (4)$$

where $\boldsymbol{\alpha}$ is a vector composed of Dirac's matrices, and Roman numerals denote variables corresponding respectively to the first and the second electrons (Kragh 1992, Roqué 1992).

While he was pondering Møller's formula and how to incorporate it into quantum electrodynamics, Dirac found a new approach to the whole theory, which he formulated in the paper "Relativistic Quantum Mechanics" dated 24 March 1932 (Dirac 1932). He transformed Møller's method quite considerably; perhaps only the initial formulation of the problem remained somewhat similar. Dirac wrote a system of two equations for two electrons:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t_1} \Psi &= (H_1 + \varepsilon_1 V(x_1, t_1)) \Psi, \\ i\hbar \frac{\partial}{\partial t_2} \Psi &= (H_2 + \varepsilon_2 V(x_2, t_2)) \Psi, \end{aligned} \qquad (5)$$

where H_1 and H_2 are Hamiltonian functions of free particles, ε_1 and ε_2 their charges, and V the field potential. Each of the two particles was characterized by its own time variable, that is why the theory later came to be called "the many-times theory." The two equations were connected through the field potential V and the total wave function Ψ, which were common to both equations. To solve them, Dirac put $t_1 = t_2 = t$ and added the equations to obtain

$$(i\hbar \frac{\partial}{\partial t} - H_1 - H_2 - \varepsilon_1 V(x_1,t) - \varepsilon_2 V(x_2,t))\,\Psi = 0. \tag{6}$$

He solved (6) by the method of successive approximations, treating the last two terms in parentheses as perturbation.

Two peculiar features of Dirac's theory need to be mentioned. First, differing from his 1927 theory of radiation and from the Heisenberg–Pauli theory, only the Hamiltonian functions for two particles were present, while the third term, the Hamiltonian function of the field itself, was lacking. Secondly, no direct Coulomb interaction between particles was introduced in the theory; electrons interacted only through the field potential V. For the latter, Dirac added the third equation, which corresponded to the free field without charges:

$$\Delta V - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} = 0, \tag{7}$$

where V was quantized in the usual manner. Therefore, in Dirac's theory, no static potential was postulated, the field consisted only of radiation.

To justify this unusual formulation of the problem, Dirac devoted more than a half of his paper to philosophical speculations, which was in general very uncharacteristic of him. He tried to imitate the discourse of the Copenhagen school, referring to the principles of correspondence and observability, but, in my view, did not do it very convincingly. This mode of thinking was alien to him; his philosophical argumentation looks self-contradictory and produces an impression of being developed *post factum*. Without presenting it at length, I will mention, as an example, that in order to justify his special treatment of the field, Dirac argued that the field plays a special role in the very process of observation and "we cannot therefore suppose the field to be a dynamical system on the same footing as the particles and thus something to be observed in the same way as the particles. The field should appear in the theory as something more elementary and fundamental" (Dirac 1932, p. 454).

Dirac presented his new theory as resulting out of a stricter compliance with the correspondence principle. For the point of departure, he chose the classical picture of electrons interacting with each other by means of absorbing and emitting waves of radiation. The elementary quantum process of this theory is thought to be a quantum jump "from the field of ingoing waves to the field of outgoing waves." He drew a parallel between the matrix element of this process with matrix elements introduced by Heisenberg in 1925 and, more generally, between the situation in quantum electrodynamics in 1932 and the situation in the old quantum theory just prior to the creation of quantum mechanics, making explicit an analogy between his new proposal and Heisenberg's revolutionary paper of 1925.

The reason for such inflated claims was, most probably, the astonishing result of his calculations. Even though they could be viewed as preliminary, because Dirac considered only the simple case of non-relativistic particles moving in a one-dimensional space and interacting through scalar waves, the second-order approximation resulted in an equation for the Ψ that corresponded to the static force of attraction between the two electrons. This inspired high hopes in him that it would be possible, in the three-dimensional case, to derive the Coulomb field out of the picture of particles exchanging radiation waves. Dirac's astonishment and enthusiasm would not have been that great, had he known that it was possible to represent the Coulomb potential with the help of waves even in classical electrodynamics, and that this method had already been applied in quantum electrodynamics by Fermi in 1930. Fermi's paper, however, had been published in Italian and was still little known when Dirac was developing his theory. Fermi obtained the Coulomb potential from waves corresponding to the scalar potential φ and the longitudinal component of the vector potential $\mathbf{A}$ of the electromagnetic field (Fermi 1930). Dirac was apparently quite surprised when he discovered this possibility by himself and hoped to solve on this basis the problems plaguing quantum electrodynamics.

His paper, however, was only a sketch of a possible theory. A realistic quantum electrodynamics would have to be constructed in a three-dimensional space, with electromagnetic waves instead of the simple scalar potential, and with particles described by Dirac's relativistic equation for the electron. In two months after the publication of Dirac's proposal, Fock and Boris Podolsky tried to meet these requirements. Podolsky, an American, was then working in Kharkov at the Ukrainian Institute of Physics and Technology. Fock came to visit there from Leningrad, and in June 1932 they co-authored two papers on the further development of Dirac's new theory (Fock and Podolsky 1932a, 1932b). The first one extended the treatment to the three-dimensional case and managed to derive the Coulomb potential with the correct sign from scalar waves: two particles with the same electric charge repelled one another, while in Dirac's one-dimensional theory they attracted one another. In the second paper, Fock and Podolsky added relativistic wave equations for electrons and the electromagnetic, rather than scalar field.

In the quantization of the electromagnetic field, they encountered the same mathematical difficulty as Heisenberg and Pauli, and as Fermi before them: the field described by the Maxwell equations could not be quantized by the usual canonical method (Heisenberg and Pauli 1929, Fermi 1929, 1930). Their version of the solution suggested quantizing a more general field and then imposing an additional constraint in order to satisfy Maxwell's equations. This additional condition was understood as a restriction on the wave function Ψ rather than on the operators of the electromagnetic field itself. This offered the third method of the quantization of electromagnetic field, which was different from, but also had something in common with, the two earlier ones (Fock and Podolsky were aware of the Heisenberg–Pauli method but had only a vague notion of Fermi's ideas). The specific forms of the generalized field and of the additional condition could vary. In their June 1932 paper, Fock and Podolsky were still unable to solve all the remaining mathematical problems, and the issue of how to bring the method of field quantization to complete consistency would still be discussed in their subsequent correspondence with each other and with Dirac.[21]

Another big problem, the relativistic description of particles, was not brought to completion either. Fock hoped to derive the retarding interaction of two electrons in the approximation to the order of $(v/c)^2$. The desired result had already existed in a less rigorous treatment by Gregory Breit (Breit 1929). Fock's calculation did not agree with the Breit formula and looked unsatisfactory: apart from the plausible terms of the order $(v/c)^2$ it also included incomprehensible imaginary terms of the order v/c. In September, Podolsky corrected the derivation, and the result then agreed with the Møller formula, rather than with the approximate Breit formula. (Podolsky and Fock 1932).[22] The circle was thus completed: the formula which had given the initial impulse to the new theoretical proposal received through it a solid justification.

While the new theory was being developed mathematically, its revolutionary value was called in question. In April 1932, Dirac presented his first proposal, which was still in press, at a conference in Copenhagen. Rosenfeld also attended the conference; he waited until Dirac's paper came out and published a critique, proving that the new theory and the old one, by Heisenberg and Pauli, were mathematically equivalent, the implication being that both were equally inadequate. Despite conspicuously different basic assumptions and equations, their formalisms could be translated into one another by a canonical transformation (Rosenfeld 1932).[23]

Despite Rosenfeld's finding, Dirac, Fock, and Podolsky proceeded, by correspondence, to develop the theory further, and in September the three met together in Leningrad at a conference on the theory of metals. Dirac and Fock presented there reports on their latest results, which agreed and overlapped to a great extent. After a vacation in the Crimea, which he spent together with Piotr Kapitza, Dirac stopped in Kharkov in early October on his way back and promised Ivanenko to contribute a paper to the newly launched Soviet journal, *Physikalische Zeitschrift der Sowjetunion.* As it turned out, Podolsky was already writing a paper on the further development of the new theory, and they finally agreed upon the publication of a joint paper by three authors. Its text was written mainly by Podolsky in Kharkov, altered and approved through extensive correspondence with Fock in Leningrad and Dirac in Cambridge, and finally resulting in the famous Dirac–Fock–Podolsky theory, dated 25 October 1932 (Dirac, Fock, and Podolsky 1932).[24]

Dirac's proposal was carried there to a completion. Main improvements over the previous Fock–Podolsky papers belonged to Dirac. He gave a simplified (compared to Rosenfeld's) proof that the new and the old quantum electrodynamics were mathematically equivalent and corrected the additional condition in the method of the quantization of electromagnetic field, which resolved remaining contradictions.[25]

Eventually, all formulations of quantum electrodynamics (Heisenberg–Pauli, Fermi, Heisenberg (1931), and Dirac–Fock–Podolsky (1932)) proved to be equivalent representations of the same theory, despite motivations to find something radically new. Although the new quantum revolution did not happen, the Dirac–Fock–Podolsky version was in several respects better than the older one by Heisenberg and Pauli. Its relativistic invariance was explicit, due to the introduction of separate time variables for each of the particles, and did not have to be proved in a complicated way. It included the so-called interaction representation, which was more convenient for calculations in most cases. In the second edition of *The Principles of Quantum Mechanics* (1935), Dirac presented quantum electrodynamics on the basis of the Dirac–Fock–Podolsky paper, and later, in the 1940s, it would play an important role in the covariant formulation of quantum electrodynamics (Tomonaga 1973, Schweber 1994, p. 277).

9. 1933: New Times

The review written by Pauli in 1932 reflects the situation of crisis in quantum electrodynamics (Pauli 1933). In his judgement, only isolated fragments of relativistic quantum theory were reliable, while attempts at unifying them into a consistent general scheme failed. Pauli recognized Dirac's theory of radiation and the Dirac equation for the electron but rejected or viewed very skeptically second quantization, his own (with Heisenberg) version of quantum electrodynamics, and the hole theory. In 1925, Pauli wrote a similar critical review of the old quantum theory (Pauli 1926), but it became outdated already in press due to the appearance of the first papers on quantum mechanics. The situation practically repeated itself with the review of 1932, this time because of the discovery of the positron.

The first announcement of a new positively charged light particle by Carl Anderson in September 1932 (Anderson 1932) was not noticed by many. The recognition came in early 1933 after an almost simultaneous publication of Anderson's second paper and the paper by Patrick Blackett and Giuseppe Occhialini (Anderson 1933, Blackett and Occhialini 1933). The view that the new particle is nothing else but the antielectron predicted by Dirac became accepted very quickly and signified the beginning of a new stage in the development of quantum electrodynamics (de Maria and Russo 1985, Roqué 1997).

A number of difficulties were solved. Combined efforts by Fermi, Heisenberg, Pauli, Dirac, Fock, and Podolsky delivered satisfactory methods of the quantization of the electromagnetic field. The discovery of the positron strengthened the credibility of the Dirac electron and the hole theory. Divergent calculations in the second order of the perturbation theory were the most serious among the remaining problems. By applying more or less artificial rules, physicists were learning how to subtract infinite values from the results and to leave only sensible finite terms that could be compared with experiment. Entirely consistent rules of dealing with this problem would not be developed, however, until the late 1940s (Schweber 1994). In the meantime, the crisis in quantum electrodynamics did not disappear entirely but became more of a chronic disease. There was still a lot of dissatisfaction with the state of the theory and a number of further attempts were made to suggest some fundamental changes in its very foundations, among the most prominent ones, by Jordan in 1933–1934, by Born and Leopold Infeld in 1934–35, by Dirac in 1936 and 1938, and a few others. The ultimate remedy, however, was as evasive as before, and a more pragmatic approach towards using the theory despite its apparent shortcomings was gradually gaining in popularity.

The discovery of the positron changed the relationship between the theory of quantum electrodynamics and experiment. Up to that point, quantum electrodynamics had hardly produced anything new for experimental physicists. Its accomplishments consisted mainly of developing a new, more rigorous and consistent foundation and justification for the existing stock of experiments and empirical formulas, which were already explained or derived with the help of earlier, less fundamental approaches. Antielectron became its first major independent prediction and it opened up a whole vast area of new phenomena, in which quantum electrodynamics had obvious advantages over earlier theories, not only logical advantages but also heuristic ones.

Despite internal imperfections it became a working theory and was providing calculations of new effects that were, within a certain area, in reasonable agreement with the growing amount of experimental data (Heitler 1936). The main line of development, apart from calculating particular effects, was represented by developing subtraction methods to handle the divergences, infinite values which continued to appear in various calculations (Rueger 1992, Brown 1993). Meanwhile, a generational change occurred, and earlier leaders, including Dirac, did not play such a crucial role any longer. Dirac's two last papers of fundamental importance for quantum electrodynamics were dated 1933.

The first one was actually written in late 1932 (Dirac 1933). Dirac returned there to his favorite idea that the Hamiltonian formalism, in which time plays a special role, is not suitable for a fundamental relativistic theory. He tried the Lagrangian dynamics instead, showing how it could be taken over into the quantum theory. Although the Lagrangian formalism has not prevailed over the Hamiltonian one in quantum field theory, Dirac's paper had a long-term consequence: in 1948 Richard Feynman transformed its approach into his new formulation of quantum mechanics on the basis of path integrals (Schweber 1994, p. 390).

In the report presented at the Leningrad conference on nuclear physics in September and at the 7th Solvay Congress in Brussels in October 1933, Dirac put forward the idea of vacuum polarization. Formulated in the terms of his hole theory, a charged particle causes displacements in the sea of negative energy electrons, which produce an observable effect similar to screening: the electric charge of the particle appears smaller than it really is (Dirac 1934a,b). While developing this idea, Dirac suggested one of the first mathematical methods of subtraction, in which a finite result of calculations is obtained as a difference between two infinite values (Dirac 1934c). Although he, thus, can be considered as an early forerunner of the renormalization method, he never regarded it as a final solution to the difficulties of quantum electrodynamics. He remained skeptical of renormalization even after it had developed into a consistent formalism, valid in all approximations of the perturbation theory. His disappointment with quantum electrodynamics, which started in the early thirties, never totally disappeared (see Kragh 1990, ch.8).

10. Conclusions

Dirac, as we saw, made great contributions to the development of practically all basic concepts and methods of quantum electrodynamics.[26] The nature of his influence, however, varied from case to case. In some cases, he created foundations of broadly accepted methods (quantization of the electromagnetic radiation, the Dirac equation, the subtraction technique), in others, he had put forward an important initial idea but gave it up later (second quantization). He developed some other approaches virtually alone or with very few supporters (the corpuscular theory of radiation, the hole theory, the prediction of the positron, the many-times theory), ignoring strong criticisms. Some of his ideas were picked up and developed further by others only years later (the Lagrangian method in quantum theory, the monopole).

Although his personal contribution to the early quantum electrodynamics is larger than anybody else's, he did not dominate the theory as much as, say, Einstein did in the case of general relativity. And the development of the theory often did not live up to his expectations either. Metaphorically speaking, his works formed the tangent vector to the trajectory, but there also always existed normal acceleration, which deflected the trajectory from the direction in which Dirac was trying to move it.

Remarks scattered among Dirac's various papers allow a reconstruction of his ideal of a theory that corresponded to his physical worldview and aesthetic tastes. Of the two related but different tasks, the relativistic generalization of quantum mechanics and the quantum treatment of electrodynamical processes, he clearly considered the first one as more fundamental. One can more properly call his preferred approach "relativistic quantum mechanics" rather than "quantum electrodynamics," and this is reflected in the titles of his papers. Relativistic invariance was, for him, the basic requirement which had to be satisfied explicitly and directly, rather than through corrections or approximations. The way to achieve this, as he repeatedly stated, was to treat time and space variables alike in the fundamental equations. This condition, however, was very hard to meet, and he himself often had to put up with relativistic modifications of the Hamiltonian formalism.

In questions of physical worldview, in his attitude towards the wave-particle dilemma, which occupied the minds of many of his contemporaries, Dirac was more inclined to view both matter and radiation in corpuscular terms. This is particularly apparent in his earlier papers and is represented by his tendency to understand radiation as an ensemble of photons, by his efforts to develop a relativistic mechanics of a single particle, and by his model of the sea of negative electrons. Even though he used mathematical formalisms of the wave theory and of quantized waves, in his physical interpretation and in discussions of obtained results, Dirac predominantly relied on corpuscular concepts.

His style was to attack fundamental issues while leaving aside calculations of concrete effects and applied problems. Dirac disliked phenomenological and hybrid theories and preferred to ground his work in some fundamental principle, such as relativistic invariance. Since fundamental requirements were difficult to satisfy all at once in one comprehensive theory, he dealt with them separately, one by one, in his various papers. Physical principles were especially important for him during the formulation of the problem and in the interpretation of final result, while in intermediate calculations he had a complete faith in mathematical formalism, even when formulas were counterintuitive.

Dirac’s disagreements with his German colleagues were increasing throughout the period discussed in this paper. They often used and developed further his mathematical approaches, but did not share his physical interpretations and preferences. In particular, Dirac’s corpuscular quantum mechanical theory of radiation and his hole theory had very few supporters. His somewhat isolated status within the field certainly contributed to his critical outlook.[27]

NOTES

[1] See the bibliography compiled by R. H. Dalitz and R. Peierls (1986) and Dirac (1995) for the papers of his last years with their titles speaking for themselves: “Does renormalization make sense?”, “The requirements of fundamental physical theory,” “The inadequacies of quantum field theory.”

[2] The map of alternative approaches within the emerging quantum mechanics and of their relative popularity is given in Kojevnikov and Novik (1989). Dirac’s fundamental contributions to quantum mechanics of that year are discussed in detail in Mehra and Rechenberg (1982), Kragh (1990), Darrigol (1992).

[3] Up to 1926, the history is covered in detail in Stuewer (1975).

[4] On the algebra of *q*-numbers, see Mehra and Rechenberg (1982) and Darrigol (1992).

[5] This was the so-called transformation theory, which later became the basis of Dirac’s book *The Principles of Quantum Mechanics* (1930b). In Copenhagen, Bohr and Heisenberg still felt the need to understand and to interpret the theory and

accomplished this in 1927 with the help of “uncertainty” and “complementarity.” For Dirac, however, quantum mechanics was completed even before these developments.

[6] He calculated coefficients for spontaneous and stimulated radiation in (Dirac 1927c) and the Kramers–Heisenberg dispersion formula in his next paper (Dirac 1927d). For a detailed analysis of Dirac’s radiation theory see Jost (1972), Darrigol (1984), Kojevnikov (1988a).

[7] On the history of early quantum electrodynamics see Cini (1982), Miller (1994), Schweber (1994, Ch.1), and Schweber (1995).

[8] The reconstruction in Kojevnikov (1988a) actually shows that second quantization preceded the transformation theory and was initially connected with Dirac’s earlier paper (Dirac 1926b). The motivation behind the quantization of the wave function was to make the value of the expression for the number of atoms in an excited state $N_r = a^*_r a_r$ an integer number. The transformation theory made it possible to derive the consequences of this proposal and to show that it leads to the description of a Bose ensemble, but it also destroyed the initial reason for it, because $a^*_r a_r$ began to be understood as probabilities and did not have to be integer. Therefore, Dirac did not mention this initial motivation when he published the idea of the second quantization in his 1927 theory of radiation.

[9] For the detailed history of the Dirac equation, see Kragh (1981b), Moyer (1981a), Kragh (1990, Ch. 3).

[10] In fact, the Klein–Gordon equation (2) can also be formally rewritten as a differential equation of the first order in time (see, e.g., Akhiezer and Berestetskiy 1981, pp. 9–10). But the main objection to Dirac’s argument came with the understanding that relativistic quantum theory is an essentially many-body theory and that its strict formulation as a one particle theory is impossible (see Pauli and Weisskopf 1934).

[11] On the program of quantized waves and wave-particle duality, see Darrigol (1986), Kojevnikov (1990a) and in this volume, Schweber (1994, pp. 33–44).

[12] Pauli to Kramers, 7 February 1928, Pauli to Dirac, 17 February 1928 (Pauli 1979, pp. 432–5).

[13] Pauli to Klein, 18 February 1929 (Pauli 1979, p. 488).

[14] More on the 1930 crisis in quantum electrodynamics in Kojevnikov (1988b, pp. 113–116), Rueger (1992).

[15] For the text of Dirac’s letter to Bohr see Moyer (1981b), for Dirac’s Paris lectures see Dirac (1931b). More on the hole theory in: Moyer (1981b), Kobzarev (1990), Kragh (1990), Dirac and Tamm (1993).

[16] Weyl (1932, Ch. 4, §6). Pauli also came to the conclusion that the masses of the electron and the hole must be equal (Tamm to Dirac, 13 September 1930, in Dirac and Tamm 1993)**.**

[17] Pauli to Meitner and others, 4 December 1930; Pauli to Klein, 12 December 1930 (Pauli 1985, pp. 39–47).

[18] For the history of the monopole, see Kragh (1981a), Krivonos (1986).

[19] Pauli to Peierls, 29 September 1931 (Pauli 1985, pp. 93–94).

[20] "Your theoretical prediction about the existence of the anti-electron . . . seemed so extravagant and totally new that you yourself dared not to cling to it and preferred rather to abandon the theory" (Tamm to Dirac, 5 June 1933, in Dirac and Tamm 1993). Tamm must have known this firsthand, because in the spring of 1931 he was in Cambridge and worked with Dirac on the theory of the monopole.

[21] Fock to Dirac, 7 July; Dirac to Fock, 19 July; Fock to Dirac, 16 October; Dirac to Podolsky, 2 November 1932 (Dirac and Fock 1990; Fock's Personal Papers in the Archive of the Russian Academy of Sciences, St. Petersburg Branch, # 1034)

[22] See also Podolsky to Fock, 26 September and 1 October 1932 (Dirac and Fock 1990). The same year, the Breit and Møller formulas were derived in Bethe and Fermi (1932) on the basis of Fermi's formulation of quantum electrodynamics.

[23] The two formulations correspond to two different representations of quantum theory, respectively the Schrödinger and the interaction representations. The characteristic feature of the interaction representation is that the equation of motion for the field operators has the form of the equation of motion for the free field, as if there are no charges (see Akhiezer and Berestetskiy 1981, p. 132). This is just the case in Dirac's equation (Eq. 7). Dirac's 1932 theory, however, uses the interaction representation only for the electromagnetic field, but not for operators describing particles.

[24] Podolsky to Fock, 9 October; Fock to Dirac, 16 October; Dirac to Podolsky, 2 November; Podolsky to Fock, 10 November; Dirac to Fock, 11 November; Podolsky to Dirac, 16 November; Podolsky to Fock, 24 November 1932 (Dirac and Fock 1990; Fock's Personal Papers in the Archive of the Russian Academy of Sciences, St. Petersburg Branch, # 1034).

[25] There was a debate between Podolsky and Dirac about the correct formulation of the equivalence proof (Dirac to Podolsky, 2 November; Podolsky to Dirac, 16 November; Podolsky to Fock, 24 November 1932). In addition to §1 of the joint paper, which belongs mostly to Dirac, Podolsky wrote §§5 and 7. Fock proved the equivalence of the additional condition in the field quantization with that given by Heisenberg and Pauli (Fock to Dirac, 16 October 1932, §6 of the joint paper).

[26] See also an appraisal in Schweber (1994, pp. 2, 70–72).

[27] The paper was originally published in Russian (Kojevnikov 1990b) in a volume on Paul Dirac which I edited together with Boris Valentinovich Medvedev and which developed from a 1986 conference held at the Institute for History of Science and Technology in Moscow. Since then, a number of important publications have appeared, including Helge Kragh's biography of Dirac, Dirac's *Collected Papers*, Sam Schweber's and others' works on the history of quantum electrodynamics, which I added to the list of references.

References

Anderson, C. A. (1932). "The Apparent Existence of Easily Deflectable Positives." *Science* 76: 238–239.

—— (1933). "The Positive Electron." *Physical Review* 43: 491–494.

Akhiezer, A. I., and Berestetskiy, V. B. (1981). *Kvantovaya electrodinamika*. 3rd ed. [*Quantum Electrodynamics*] Moscow: Nauka.

Blackett, P. M. S., and Occhialini, G. P. S. (1933). "Some Photographs of the Tracks of Penetrating Radiation." *Royal Society of London*. *Proceedings* 139: 699–726.

Bethe, H., and Fermi, E. (1932). "Über Wechselwirkung von zwei Electronen." *Zeitschrift für Physik* 77: 296–306.

Born, Max (1926). "Das Adiabatenprinzip in der Quantenmechanik." *Zeitschrift für Physik* 40: 167–192.

Born, M., Heisenberg, W., and Jordan, P. (1926). "Zur Quantenmechanik, II." *Zeitschrift für Physik* 35: 557–615.

Breit, G. (1929). "The Effect of Retardation on the Interaction of Two Electrons." *Physical Review* 34: 553–573.

Bromberg, Joan (1976). "The Concept of Particle Creation before and after Quantum Mechanics." *Historical Studies in Physical Sciences* 7: 161–191.

Brown, Laurie, ed. (1993). *Renormalization: From Lorentz to Landau (and Beyond).* New York: Springer.

Cini, M. (1982) "Cultural Traditions and Environmental Factors in the Development of Quantum Electrodynamics (1925–1933)." *Fundamenta Scientiae* 3: 229–253.

Dalitz, R. H., and Peierls, R. (1986). "P. A. M. Dirac. 1902–1984." *Royal Society of London*. *Biographical Memoirs* 32: 139–185.

Darrigol, Olivier (1984). "La genèse du concept de champ quantique." *Annales de Physique Française* 9: 433–501.

––– (1986). "The Origin of Quantized Matter Waves." *Historical Studies in Physical and Biological Sciences* 16: 197–253.

––– (1992). *From C-numbers to Q-numbers: The Classical Analogy in the History of Quantum Theory*. Berkeley: University of California Press.

Darwin, C. G. (1927). "The Electron as a Vector Wave." *Royal Society of London. Proceedings* 116: 227–233.

de Maria, M., and Russo, A. (1985). "The Discovery of the Positron." *Rivista di Storia della Scienza* 2: 237–286.

Dirac, P. A. M. (1926a). "Relativity Quantum Mechanics with an Application to Compton Scattering." *Royal Society of London*. *Proceedings* 111: 405–423.

––– (1926b). "On the Theory of Quantum Mechanics." *Royal Society of London*. *Proceedings* 112: 661–677.

––– (1927a). "The Compton Effect in Wave Mechanics." *Cambridge Philosophical Society*. *Proceedings* 23: 500–507.

––– (1927b). "The Physical Interpretation of the Quantum Dynamics." *Royal Society of London*. *Proceedings* 113: 621–641.

––– (1927c). "The Quantum Theory of the Emission and Absorption of Radiation." *Royal Society of London. Proceedings* 114: 243–256.

––– (1927d). "The Quantum Theory of Dispersion." *Royal Society of London*. *Proceedings* 114: 710–728.

––– (1928a). "The Quantum Theory of the Electron." *Royal Society of London*. *Proceedings* 117: 610–624.

––– (1928b). "The Quantum Theory of the Electron, Part II." *Royal Society of London*. *Proceedings* 118: 351–361.

––– (1928c). "Zur Quantentheorie des Elektrons." In *Quantentheorie und Chemie*. Hans Falkenhagen and S. Hirzel, eds. Leipzig: S. Hirzel, pp. 85–94.

––– (1929). "Quantum Mechanics of Many-Electron Systems." *Royal Society of London*. *Proceedings* 123: 714–733.

––– (1930a). "A Theory of Electrons and Protons." *Royal Society of London*. *Proceedings* 126: 360–365.

––– (1930b). *The Principles of Quantum Mechanics*. Oxford: Clarendon Press.

––– (1930c). "On the Annihilation of Electrons and Protons." *Cambridge Philosophical Society*. *Proceedings* 26: 376–385.

––– (1931a). "Quantized Singularities in the Electromagnetic Field." *Royal Society of London*. *Proceedings* 133: 60–72.

––– (1931b). "Quelquez problèmes de mécanique quantique." *Institut Henri Poincaré*. *Annales* 1: 357–400.

––– (1932). "Relativistic Quantum Mechanics." *Royal Society of London*. *Proceedings* 136: 453–464.

––– (1933). "The Lagrangian in Quantum Mechanics." *Physikalische Zeitschrift der Sowjetunion* 3: 64–72.

––– (1934a). "Teoriya pozitrona" ["A Theory of the Positron"]. In *Atomnoye yadro* [*The Atomic Nucleus*]. M. P. Bronstein, V. M. Dukel'skiy, D. D. Ivanenko, and Yu. B. Khariton, eds. Leningrad and Moscow: Gostekhteorizdat, pp. 129–143.

––– (1934b). "Theorie du positron." In *Structure et propriétés des noyaux atomiques*. Rapports et discussions du septième conseil de physique tenu à Bruxelles du 22 au 29 Octobre 1933. Paris: Gauthier-Villars, pp. 203–230.

––– (1934c). "Discussion of the Infinite Distribution of Electrons in the Theory of the Positron." *Cambridge Philosophical Society*. *Proceedings* 30: 150–163.

––– (1935). *The Principles of Quantum Mechanics*. 2nd ed. Oxford: Clarendon Press.

––– (1959). "Sovremennoe sostoyanie reliativistskoy teorii elektrona." ["Contemporary State of the Relativistic Theory of the Electron."] *Institut istorii estestvoznaniya i tekhniki. Trudy* [*Institute for the History of Science and Technology*. *Transactions*] 22: 32–33.

––– (1983). "The Origin of Quantum Field Theory." In *The Birth of Particle Physics*. L. M. Brown and L. Hoddeson, eds. Cambridge and New York: Cambridge University Press, pp. 39–55.

––– (1995). *The Collected Works of P. A. M. Dirac, 1924–1948*. R. H. Dalitz, ed. Cambridge and New York: Cambridge University Press.

Dirac, P. A. M., and Fock, V. A. (1990). "Perepiska Diraka s V. A. Fokom. Materialy k istorii stat'i Diraka–Foka–Podol'skogo." ["Dirac's Correspondence with V. A. Fock. Documents on the History of the Dirac–Fock–

Podolsky Theory"] A. B. Kozhevnikov, V. Ya. Frenkel, N. I. Nevskaya, and T. I. Efremidze, eds. In: Medvedev and Kozhevnikov 1990, pp. 177–197.

Dirac, P. A. M., Fock, V. A., and Podolsky, B. (1932). "On Quantum Electrodynamics." *Physikalische Zeitschrift der Sowjetunion* 2: 468–479.

Dirac, P. A. M., and Tamm, I. E. (1993–1996). *Paul Dirac and Igor Tamm. Correspondence*. Part 1: 1928–1933; part 2: 1933–1936. A. B. Kojevnikov, ed. Munich: Max-Planck-Institut für Physik. Preprints 93–80 and 96–40.

Fermi, E. (1929). "Sopra l'elettrodinamica quantistica." *Accademia dei Lincei. Rendiconti* 9: 881–887.

––– (1930). "Sopra l'elettrodinamica quantistica." *Accademia dei Lincei. Rendiconti* 10: 431–435.

Fock, V. A. (1932). "Konfigurationsraum und zweite Quantelung." *Zeitschrift für Physik* 75: 622–647; 76: 852.

Fock, V. A., and Podolsky, B. (1932a). "Zur Diracschen Quantenelektrodynamik." *Physikalische Zeitschrift der Sowjetunion* 1: 798–800.

––– (1932b). "On the Quantization of Electromagnetic Waves and the Interaction of Charges in Dirac's Theory." *Physikalische Zeitschrift der Sowjetunion* 1: 801–817.

Gordon, W. (1926). "Der Comptoneffekt nach der Schrödingerschen Theorie." *Zeitschrift für Physik* 40: 117–133.

Heisenberg, W. (1931). "Bemerkungen zur Strahlungstheorie." *Annalen der Physik* 9: 338–346.

Heisenberg, W., and Pauli, W. (1929). "Zur Quantendynamik der Wellenfelder." *Zeitschrift für Physik* 56: 1–61.

––– (1930). "Zur Quantentheorie der Wellenfelder. II." *Zeitschrift für Physik* 59: 168–190.

Heitler, W. (1936). *The Quantum Theory of Radiation.* Oxford: Clarendon Press.

Jordan, P. (1927a). "Uber eine neue Begründung der Quantenmechanik." *Zeitschrift für Physik* 40: 809–838.

––– (1927b). "Über eine neue Begründung der Quantenmechanik." *Gesellschaft der Wissenschaften zu Göttingen. Nachrichten*: 161–169.

––– (1927c). "Zur Quantenmechanik der Gasentartung." *Zeitschrift für Physik* 44: 473–480.

––– (1928). "Die Lichtquantenhypothese. Entwicklung und gegenwärtige Stand." *Ergebnisse der Exakten Wissenschaften* 7: 158–208.

Jordan, P., and Klein, O. (1927). "Zum Mehrkörperproblem der Quantentheorie." *Zeitschrift für Physik* 45: 751–765.

Jordan, P., and Pauli, W. (1928). "Zur Quantenelektrodynamik ladungsfreier Felder." *Zeitschrift für Physik* 47: 151–173.

Jordan, P., and Wigner, E. (1928). "Uber das Paulische Äquivalenzverbot." *Zeitschrift für Physik* 47: 631–651.

Jost, Res (1972). "Foundations of Quantum Field Theory." In *Aspects of Quantum Theory*. A. Salam and E. P. Wigner, eds. Cambridge: Cambridge University Press, pp. 61–77.

Klein, O. (1927). “Elektrodynamik und Wellenmechanik vom Standpunkt des Korrespondenzprinzips.” *Zeitschrift für Physik* 41: 407–442.

—– (1929). “Die Reflexion von Elektronen an einem Potentialsprung nach der relativistischen Dynamik von Dirac.” *Zeitschrift für Physik* 53: 157–165.

Klein, O., and Nishina, Y. (1929). “Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac.” *Zeitschrift für Physik* 52: 853–868.

Kobzarev, I. Yu. (1990). “K istorii positrona.” [“On the History of the Positron”] In Medvedev and Kojevnikov 1990, pp. 21–34.

Kojevnikov, A. B. (1988a). “Dirak i kvantovaya teoriya izlucheniya.” [“Dirac and the Quantum Theory of Radiation”] In *Einshteinovskiy sbornik 1984–85*. I. Yu. Kobzarev and G. E. Gorelik , eds. Moscow: Nauka, pp. 246–270.

—– (1988b). “V. A. Fock i metod vtorichnogo kvantovaniya.” [“V. A. Fock and the Method of Second Quantization”] In: *Issledovaniya po istorii fiziki i mekhaniki, 1988*. (*Studies in the History of Physics and Mechanics, 1988*) A. T. Grigorian, ed. Moscow: Nauka, pp. 115–138.

—– (1990a). “Einshteinovskaya formula dlya fluktuatsii i korpuskulyarno-volnovoy dualizm.” [“Einstein’s Fluctuation Formula and the Wave-Particle Duality”] In *Einshteinovskiy sbornik 1986–87*. I. Yu. Kobzarev and G. E. Gorelik, eds. Moscow: Nauka, pp. 102–124.

—– (1990b). “Kvantovaya elektrodinamika Diraka.” [“Dirac’s Quantum Electrodynamics”] In Medvedev and Kojevnikov 1990, pp. 44–66.

Kojevnikov, A. B., and Novik, O. I. (1989). “Analysis of Informational Ties in Early Quantum Mechanics (1925–1927).” In: *Acta Historiae Rerum Naturalium nec non Technicarum.* Iss. 20. Prague, pp. 115–159.

Kragh, Helge (1981a). “The Concept of the Monopole. A Historical and Analytical Case Study.” *Studies in History and Philosophy of Science* 12: 141–172.

—– (1981b). “The Genesis of Dirac’s Relativistic Theory of Electrons.” *Archive for History of the Exact Sciences* 24: 31–67.

—– (1984). “Equation with Many Fathers: The Klein–Gordon Equation in 1926.” *American Journal of Physics* 52: 1024–33.

—– (1990). *Dirac. A Scientific Biography*. Cambridge: Cambridge University Press

—– (1992). “Relativistic Collisions: The Work of Christian Møller in the Early 1930’s.” *Archive for History of the Exact Sciences* 43: 299–328.

Krivonos, L. N. (1986). “Istoriya gipotezy monopolya Diraka.” [”The History of Dirac’s Monopole Hypothesis”] *NTM. Schriftenreihe für Geschichte der Naturwissenschaften, Technik und Medizin* 23: 35–45.

Medvedev, B. V., and Kojevnikov, A. B. eds. (1990). *Pol’ Dirak i fizika XX veka.* [*Paul Dirac and Twentieth-Century Physics*] Moscow: Nauka.

Mehra, Jagdish, and Rechenberg, Helmut (1982). *The Historical Development of Quantum Theory*. Vol. 4. New York: Springer-Verlag.

Miller, Arthur I. (1994). *Early Quantum Electrodynamics. A Source Book.* Cambridge: Cambridge University Press.

Møller, Christian (1931), “Über den Stoß zweier Teilchen unter Berücksichtigung der Retardation der Kräfte.” *Zeitschrift für Physik* 70: 786–795.

Moyer, Donald Franklin (1981a). “Origins of Dirac’s Electron, 1925–1928.” *American Journal of Physics* 49: 944–948.

–– (1981b). “Evaluations of Dirac’s Electron, 1928–1932.” *American Journal of Physics* 49: 1055–1062.

Oppenheimer, J.R. (1930). “On the theory of Electrons and Protons.” *Physical Review* 35: 562–563.

Pauli, W. (1926). “Quantentheorie.” *Handbuch der Physik.* Bd 23. Berlin: Springer pp. 1–278.

–– (1927). “Zur Quantenmechanik des magnetischen Elektrons.” *Zeitschrift für Physik* 43: 601–623.

–– (1933). “Die allgemeinen Prinzipien der Wellenmechanik.” *Handbuch der Physik.* Bd 24, T. 1. Berlin: Springer, pp. 82–272.

–– (1979). *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* 2 vols. Karl von Meyenn, ed. New York: Springer-Verlag.

Pauli, W., and Weisskopf, V. (1934). “Uber die Quantizierung der skalaren relativistischen Wellengleichung.” *Helvetica Physica Acta* 7: 709–731.

Podolsky, B., and Fock, V. A. (1932). “Derivation of Möller’s Formula from Dirac’s Theory.” *Physikalische Zeitschrift der Sowjetunion* 2: 275–277.

Roqué, Xavier (1992). “Møller Scattering: A Neglected Application of Early Quantum Electrodynamics.” *Archive for History of the Exact Sciences* 44: 197–264.

–– (1997) “The Manufacture of the Positron.” *Studies in History and Philosophy of Modern Physics* 28: 73–129.

Rosenfeld, Leon (1931a). “Zur Kritik der Diracschen Strahlungstheorie.” *Zeitschrift für Physik* 70: 454–462.

–– (1931b). “Zur korrespondenzmäßigen Behandlung der Linienbreite.” *Zeitschrift für Physik* 71: 273–278.

–– (1931c). “Bemerkung zur korrespondenzmäßigen Behandlung des relativistischen Mehrkörperproblems.” *Zeitschrift für Physik* 73: 253–259.

–– (1932). “Über eine mögliche Fassung des Diracschen Programms zur Quantenelektrodynamik und deren formalen Zusammenhang mit der Heisenberg–Paulischen Theorie.” *Zeitschrift für Physik* 76: 729–734.

Rueger, Alexander. (1992). “Attitudes towards Infinities: Responses to Anomalies in Quantum Electrodynamics, 1927–1947.” *Historical Studies in the Physical and Biological Sciences* 22: 309–337.

Schweber, Silvan S. (1994). *QED and the Men who Made it: Dyson, Feynman, Schwinger, and Tomonaga*. Princeton: Princeton University Press.

–– (1995). “Early Quantum Electrodynamics.” *Studies in History and Philosophy of Modern Physics* 26: 201–211.

Slater, J. C. (1927). “Radiation and Absorption on Schrödinger’s Theory.” *National Academy of Sciences of the USA. Proceedings* 13: 7–12.

Small, Henry (1983). *Physics Citation Index, 1920–1929.* Philadelphia: Institute for Scientific Information.

Stuewer, Roger (1975). *The Compton Effect: Turning Point in Physics*. New York: Science History Publications.

Tamm, I. (1930). "Über die Wechselwirkung der freier Elektronen mit der Strahlung nach der Diracschen Theorie des Elektrons und nach Quantenelektrodynamik." *Zeitschrift für Physik* 62: 545–568.

Tomonaga, S. I. (1973). "Development of Quantum Electrodynamics." In *The Physicist's Conception of Nature*. J. Mehra, ed. Dordrecht: Reidel, pp. 404–412.

Weyl, H. (1972). *The Theory of Groups and Quantum Mechanics*. H. P. Robertson, trans. New York: Dutton.